AD NO. ___________________________
DTC PROJECT NO. 8-CO-160-UXO-021
REPORT NO. ATC-9813

STANDARDIZED

UXO TECHNOLOGY DEMONSTRATION SITE

OPEN FIELD SCORING RECORD NO. 914

SITE LOCATION:
U.S. ARMY YUMA PROVING GROUND

DEMONSTRATOR:
GEOCENTERS SAIC
7 WELLS AVENUE
NEWTON, MA 02459

TECHNOLOGY TYPE/PLATFORM:
VEHICULAR SIMULTANEOUS EMI AND
MAGNETOMETER SYSTEM (VSEMS)/TOWED

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

SEPTEMBER 2008

Prepared for:
U.S. ARMY ENVIRONMENTAL COMMAND
ABERDEEN PROVING GROUND, MD 21010-5401

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **SEP 2008**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Standardized UXO Technology Demonstration Site Blind Grid Scoring Record No. 915 (Geometrics, Inc.)**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Commander U.S. Aberdeen Test Center ATTN: TEDT-AT-SLE Aberdeen Proving Ground, MD 21005-5059**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**51**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

DISPOSITION INSTRUCTIONS

Destroy this document when no longer needed. Do not return to the originator.

TEDT-AT-SLE

MEMORANDUM FOR RECORD

SUBJECT: Operations Security (OPSEC) Review of Paper/Presentation

1. The attached document entitled "Scoring Record No. 914" dated September 2008 is provided for review for public disclosure in accordance with AR 530-1 as supplemented. The document is proposed for public release via the internet.

2. I, the undersigned, am aware of the intelligence interest in open source publications and in the subject matter of the information I have reviewed for intelligence purposes. I certify that I have sufficient technical expertise in the subject matter of this document and that, to the best of my knowledge, the net benefit of this public release outweighs the potential damage to the essential secrecy of all related ATC, DTC, ATEC, Army or other DOD programs of which I am aware.

J. Stephen McClung		September 2008
NAME (Printed)	SIGNATURE	DATE

CONCURRENCE:	NAME (Printed)	SIGNATURE	DATE
Program Mgr/Customer (If not ATC owned technology)	Patrick McDonnell		9/30/08
Directorate Director	Charles Valz		10/14/08
Directorate OPSEC QC and Team Leader	William Burch		16 Sep 08
ATC OPSEC Officer/ Security Manager	Jenell Bigham		20 Oct 08
Public Affairs Specialist	Crystal Maynard		22 Oct 08
Technical Director, ATC	John R. Wallace		23 Oct 08
(Return to ATC PAO for further processing)			
DTC GO/SES	N/A		

Encl
as

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
September 2008	Final	12 through 14, 16 June 2006

4. TITLE AND SUBTITLE
STANDARDIZED UXO TECHNOLOGY DEMONSTRATION SITE OPEN FIELD SCORING RECORD NO.914 (GEOCENTERS SAIC)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
McClung, J. Stephen

5d. PROJECT NUMBER
8-CO-160-UXO-021

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Aberdeen Test Center
ATTN: TEDT-AT-SLE
Aberdeen Proving Ground, MD 21005-5059

8. PERFORMING ORGANIZATION REPORT NUMBER
ATC-9813

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Environmental Command
ATTN: IMAE-RTA
Aberdeen Proving Ground, MD 21005-5401

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Same as Item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited.

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT

This scoring record documents the efforts of Geocenters SAIC to detect and discriminate inert unexploded ordnance (UXO) utilizing the YPG Standardized UXO Technology Demonstration Site Open Field. This Scoring Record was coordinated by J. Stephen McClung and the Standardized UXO Technology Demonstration Site Scoring Committee. Organizations on the committee include the U.S. Army Corps of Engineers, the Environmental Security Technology Certification Program, the Strategic Environmental Research and Development Program, the Institute for Defense Analysis, the U.S. Army Environmental Command, and the U.S. Army Aberdeen Test Center.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER *(Include area code)*
Unclassified	Unclassified	Unclassified	SAR		

ACKNOWLEDGMENTS

Authors:

Rick Fling
Aberdeen Test Support Services (ATSS)
Sverdrup Technology, Inc.
U.S. Army Aberdeen Proving Ground (APG)

Christina McClung
Aberdeen Data Services Team (ADST)
Logistics Engineering and Information Technology Company (Log.Sec/Tri-S)
U.S. Army Aberdeen Proving Ground

Contributors:

William Burch
J. Stephen McClung
Military Environmental Technology Demonstration Center (METDC)
U.S. Army Aberdeen Test Center (ATC)
U.S. Army Aberdeen Proving Ground

Leonardo Lombardo
Aberdeen Test and Support Services
Sverdrup Technology, Inc.
U.S. Army Aberdeen Proving Ground

Patrick McDonnell
Booz Allen Hamilton (BAH)
U.S. Army Environmental Command (USAEC)
U.S. Army Aberdeen Proving Ground

TABLE OF CONTENTS

SECTION 4. TECHNICAL PERFORMANCE RESULTS

PAGE

SECTION 5. ON-SITE LABOR COSTS

SECTION 6. COMPARISON OF RESULTS TO BLIND GRID DEMONSTRATION

SECTION 7. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of munitions and explosives of concern (MEC) – i.e., unexploded ordnance (UXO) and discarded military munitions (DMM) require testing so that their performance can be characterized. To that end, Standardized Test Sites have been developed at Aberdeen Proving Ground (APG), Maryland, and U.S. Army Yuma Proving Ground (YPG), Arizona. These test sites provide a diversity of geology, climate, terrain, and weather as well as diversity in ordnance and clutter. Testing at these sites is independently administered and analyzed by the Government for the purposes of characterizing technologies, tracking performance with system development, comparing performance of different systems, and comparing performance in different environments.

The Standardized UXO Technology Demonstration Site Program is a multiagency program spearheaded by the U.S. Army Environmental Command (USAEC). The U.S. Army Aberdeen Test Center (ATC) and the U.S. Army Corps of Engineers Engineering Research and Development Center (ERDC) provide programmatic support. The program is being funded and supported by the Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP), and the Army Environmental Quality Technology Program (EQT).

1.2 SCORING OBJECTIVES

The objective in the Standardized UXO Technology Demonstration Site Program is to evaluate the detection and discrimination capabilities of a given technology under various field and soil conditions. Inert munitions and clutter items are positioned in various orientations and depths in the ground.

The evaluation objectives are as follows:

a. To determine detection and discrimination effectiveness under realistic scenarios that vary targets, geology, clutter, topography, and vegetation.

b. To determine cost, time, and manpower requirements to operate the technology.

c. To determine the demonstrator's ability to analyze survey data in a timely manner and provide prioritized "Target Lists" with associated confidence levels.

d. To provide independent site management to enable the collection of high quality, ground-truth, geo-referenced data for post-demonstration analysis.

1.2.1 Scoring Methodology

a. The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating

characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}), and those that do not correspond to any known item, termed background alarms.

b. The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the blind grid RESPONSE STAGE, the demonstrator provides the scoring committee with a target response from each and every grid square along with a noise level below which target responses are deemed insufficient to warrant further investigation. This list is generated with minimal processing and, since a value is provided for every grid square, will include signals both above and below the system noise level.

c. The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such and to reject clutter. For the blind grid DISCRIMINATION STAGE, the demonstrator provides the scoring committee with the output of the algorithms applied in the discrimination-stage processing for each grid square. The values in this list are prioritized based on the demonstrator's determination that a grid square is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For digital signal processing, priority ranking is based on algorithm output. For other discrimination approaches, priority ranking is based on human (subjective) judgment. The demonstrator also specifies the threshold in the prioritized ranking that provides optimum performance (i.e., that is expected to retain all detected ordnance and rejects the maximum amount of clutter).

d. The demonstrator is also scored on EFFICIENCY and REJECTION RATIO, which measures the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. EFFICIENCY measures the fraction of detected ordnance retained after discrimination, while the REJECTION RATIO measures the fraction of false alarms rejected. Both measures are defined relative to performance at the demonstrator-supplied level below which all responses are considered noise, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

e. Based on configuration of the ground truth at the standardized sites and the defined scoring methodology, there exists the possibility of having anomalies within overlapping halos and/or multiple anomalies within halos. In these cases, the following scoring logic is implemented:

(1) In situations where multiple anomalies exist within a single R_{halo}, the anomaly with the strongest response or highest ranking will be assigned to that particular ground truth item.

(2) For overlapping R_{halo} situations, ordnance has precedence over clutter. The anomaly with the strongest response or highest ranking that is closest to the center of a particular ground truth item gets assigned to that item. Remaining anomalies are retained until all matching is complete.

(3) Anomalies located within any R_{halo} that do not get associated with a particular ground truth item are thrown out and are not considered in the analysis.

f. All scoring factors are generated utilizing the Standardized UXO Probability and Plot Program, version 3.1.1.

1.2.2 Scoring Factors

Factors to be measured and evaluated as part of this demonstration include:

a. Response Stage ROC curves:

(1) Probability of Detection (P_d^{res}).

(2) Probability of False Positive (P_{fp}^{res}).

(3) Background Alarm Rate (BAR^{res}) or Probability of Background Alarm (P_{BA}^{res}).

b. Discrimination Stage ROC curves:

(1) Probability of Detection (P_d^{disc}).

(2) Probability of False Positive (P_{fp}^{disc}).

(3) Background Alarm Rate (BAR^{disc}) or Probability of Background Alarm (P_{BA}^{disc}).

c. Metrics:

(1) Efficiency (E).

(2) False Positive Rejection Rate (R_{fp}).

(3) Background Alarm Rejection Rate (R_{BA}).

d. Other:

(1) Probability of Detection by Size and Depth.

(2) Classification by type (i.e., 20-, 40-, 105-mm, etc.).

(3) Location accuracy.

(4) Equipment setup, calibration time, and corresponding man-hour requirements.

(5) Survey time and corresponding man-hour requirements.

(6) Reacquisition/resurvey time and man-hour requirements (if any).

(7) Downtime due to system malfunctions and maintenance requirements.

1.3 STANDARD AND NONSTANDARD INERT ORDNANCE TARGETS

The standard and nonstandard ordnance items emplaced in the test areas are listed in Table 1. Standardized targets are members of a set of specific ordnance items that have identical properties to all other items in the set (caliber, configuration, size, weight, aspect ratio, material, filler, magnetic remanence, and nomenclature). Nonstandard targets are inert ordnance items having properties that differ from those in the set of standardized targets.

TABLE 1. INERT ORDNANCE TARGETS

Standard Type	Nonstandard (NS)
20-mm Projectile M55	20-mm Projectile M55
	20-mm Projectile M97
40-mm Grenades M385	40-mm Grenades M385
40-mm Projectile MKII Bodies	40-mm Projectile M813
BDU-28 Submunition	
BLU-26 Submunition	
M42 Submunition	
57-mm Projectile APC M86	
60-mm Mortar M49A3	60-mm Mortar (JPG)
	60-mm Mortar M49
2.75-inch Rocket M230	2.75-inch Rocket M230
	2.75-inch Rocket XM229
MK 118 ROCKEYE	
81-mm Mortar M374	81-mm Mortar (JPG)
	81-mm Mortar M374
105-mm HEAT Rounds M456	
105-mm Projectile M60	105-mm Projectile M60
155-mm Projectile M483A1	155-mm Projectile M483A
	500-lb Bomb

JPG = Jefferson Proving Ground
HEAT = high-explosive antitank

SECTION 2. DEMONSTRATION

2.1 DEMONSTRATOR INFORMATION

2.1.1 Demonstrator Point of Contact (POC) and Address

POC: Mr. Rob Siegel
(617) 618-4662

Address: Geocenters SAIC
7 Wells Avenue
Newton, MA 02459

2.1.2 System Description (provided by demonstrator)

The vehicular simultaneous electromagnetic interference (EMI) and magnetometer system (VSEMS) (fig. 1) is a vehicle-towed array that simultaneously collects total field magnetometer (MAG) and electromagnetic (EM) 61 MKII data. Normally these two sensors cannot be deployed within about 30 feet of each other because the active nature of the EM61 sensor creates noise on the magnetometers, but VSEMS contains patented-applied-for electronics that interleave the two data streams, monitoring the EM61 sync pulse and waiting until the secondary fields it creates have died down before sampling the magnetometers when the EM61s are quiet. In this way, we concurrently collect high-quality EM61 and magnetometer data in a single survey pass. New to VSEMS for this fielding are a carbon fiber platform designed for survivability and minimizing unregistered sensor motion, improved system timing in the electronics and software designed to time-stamp the sensor updates as accurately as possible, and three Global Positioning System (GPS) receivers on the platform designed to position the sensors as accurately as possible. All of these modifications contribute to the goal of improving the accuracy of each position update from each sensor.

Figure 1. Demonstrator's system, VSEMS dual/towed.

Tow Vehicle	Custom-built aluminum-framed dune buggy with very low magnetic self-signature.
Sensor Platform	Custom-built fiberglass platform, reinforced with marine-grade plywood, with titanium suspension to host both magnetometers and EM61s in low-noise environment. New carbon fiber platform under development.
Magnetometers	Five Geometrics 822A aircraft quality cesium vapor total field magnetometers
Magnetometer Interface	Science Applications International Corporation (SAIC's) custom MAG Period Counter (developed under ESTCP Project No. UX-0208) that interleaves magnetometer data between EM61 pulses. Unique to SAIC and patent applied for.
Magnetometer Sampling Rate	75 Hz interleaved between EM61 pulses.
EM61 Configuration	Five Geonics EM61 MKII's (4 time gates) driving five 1 by ½-meter coils arranged with the short axis cross-track for maximum cross-track resolution.
EM61 Sampling Rate	75 Hz internal; 10 Hz serial output.
Sensor Swath	2.5 meters (EM61 coils edge-to-edge).
GPS	Trimble real-time kinematic (RTK)-equipped system for 2-cm accuracy in real time.
GPS-Magnetometer Synchronization	Magnetometers triggered by GPS 1 PPS signal, guaranteeing acquisition of correctly synchronized data (patented and unique to VSEMS).
Survey Speed	1 to 5 mph on smooth, level, vegetation-free terrain.
Surveys	Nearly 900 acres of real-world UXO and MEC sites.

2.1.3 Data Processing Description (provided by demonstrator)

Most data processing occurs in custom Linux-based software. The software internally converts the GPS data from latitude and longitude into Universal Transverse Mercator (UTM) coordinates. The GPS data were read to determine the spatial extent of the site surveyed. The software then set up a site (a grid in memory) that wholly contained the surveyed data. The position data were examined and corrected as needed. Automatic correction examined the position data for jumps that were greater than possible for low-speed vehicular data. The heading between updates was determined, and the position of the 75-Hz MAG and 10-Hz EM samples were calculated. If large jumps in the position data were encountered (e.g., jumps caused by short-term differential dropouts), the operator was asked to examine the data and hand-correct a bad point by forcing it to align with the normal survey line. The corrected navigation data were then saved with the sensor data in a new file.

The MAG data were notch-filtered to remove the 60-Hz electrical hum that is pervasive around buildings and power lines. The MAG and EM61 data were then background-leveled. Typically, a median filter is used to determine the background reading for a 5-second window, and then this background is subtracted out. The EM61 data are then latency-corrected by visually inspecting the data and adjusting the latency correction so that portions of anomalies acquired in opposite directions lined up. The data were then gridded. A linear inverse distance squared interpolation was used, with an interpolation window of ± 30 cm. This interpolation window functioned in both directions. Interpolation was performed cross-track (between the sensors spaced 1/2 meter apart) as well as along the direction of travel (between the 75-Hz MAG or 10 Hz-EM updates).

Target analysis commenced once the interpolated image was displayed. The operator adjusted a zoom-in factor and display scales (e.g., $\pm$ 250 gammas) or gray scale highlights (e.g., highlight every reading over 50 gammas as red). Anomaly analysis is accomplished with the operator selecting an area of interest (AOI) around an anomaly. The data high and low values were determined and displayed inside the AOI and were used as a seed, based on the full width at half maximum rule, for the model match. The data inside the AOI were then matched recursively to a magnetic dipole model. The results of the model match provided anomaly location and estimated depth and size. These parameters, along with optional operator comments, were logged into the site target file. This procedure worked well for isolated anomalies. Complex anomalies, caused by clusters of multiple objects or geology, required more expert operator interaction. If the magnetic dipole model failed to converge, or converged on an impossible or unphysical result, the operator could log the location based on the full width at half maximum rule. In this case, no estimate of depth or size could be produced. The operator could also pinpoint a target's location with the mouse. Again, no depth or size estimates were available for such targets. A sequential number was assigned to each target. For large, complex, extended areas of contamination, an operator could create a perimeter landmark file that logged pinpoint locations selected by the operator. MAG and pulsed EM data were analyzed simultaneously by running two copies of the VSEMS workstation software and linking them together so that panning and scrolling in one pans and scrolls in the other, and drawing an area of interest in one draws the same area of interest in the other.

When the operator completed the analysis, the target report was output in a format suitable for importation into other tools such as Excel and Word. The target report contained targets from both the MAG and EM61 data. The coordinates used in the report file were transformed to the required system (e.g., State Grid Plane or UTM).

2.1.4 Data Submission Format

Data were submitted for scoring in accordance with data submission protocols outlined in the Standardized UXO Technology Demonstration Site Handbook. These submitted data are not included in this report in order to protect ground truth information.

2.1.5 Demonstrator Quality Assurance (QA) and Quality Control (QC) (provided by demonstrator)

QC. Base GPS was set up over a known control point. Sensors were warmed up for 5 minutes prior to data collection. An object was passed in front of each sensor, and the response on the vehicle computer was examined to verify that each sensor was operating and connected to the correct channel. Prior to coming to the site, cable shake issues and approximate EM61 latency issues were resolved. When required, additional Corps of Engineers - Huntsville Center (CEHNC)-mandated QC static tests, shake tests, and six-line tests were performed on-site.

QA. Geocenters SAIC operators understand geophysics and sensors and know when things are working and when they are not. Numerical outputs from the sensors and the GPS are displayed at all times in the vehicle, and these values were examined at the start and finish of every survey line. A small test set of data on the site were acquired, processed, and examined

to verify that there would be no surprises. An automated data quality program examines the data and reports out-of-range MAG readings and bad (non-differential) position readings. This gave a quick and convenient benchmark on out-of-range data that may be indicative of navigation or sensor errors. Vehicle data were downloaded at lunch, and the data were examined to ensure that no degradation had occurred since morning. Data were downloaded again at the end of the day. Coordinates of the site and grids were overlaid on the site over data to verify that data were being correctly positioned. All GPS data were examined and hand-corrected when necessary if the radio link between base and rover was interrupted. All sensor data were examined by hand. All magnetometer and EM61 data were background-leveled. Latency-correcting the MAG data was not necessary since Geocenters SAIC's hardware is designed to trigger the magnetometers using the GPS' 1 PPS, which guarantees latency-free data. The EM61 data were latency corrected in an industry-standard fashion, lining up halves of anomalies acquired in opposite direction.

2.1.6 Additional Records

The following record(s) by this vendor can be accessed via the Internet as Microsoft Word documents at www.uxotestsites.org. The Blind Grid Counterpart to this report is Scoring Record No. 792.

2.2 YPG SITE INFORMATION

2.2.1 Location

YPG is located adjacent to the Colorado River in the Sonoran Desert. The UXO Standardized Test Site is located south of Pole Line Road and east of the Countermine Testing and Training Range. The open field range, calibration grid, blind grid, mogul area, and desert extreme area comprise the 350- by 500-meter general test site area. The open field site is the largest of the test sites and measures approximately 200 by 350 meters. To the east of the open field range are the calibration and blind test grids that measure 30 by 40 meters and 40 by 40 meters, respectively. South of the open field is the 135- by 80-meter mogul area consisting of a sequence of man-made depressions. The desert extreme area is located southeast of the open field site and has dimensions of 50 by 100 meters. The desert extreme area, covered with desert-type vegetation, is used to test the performance of different sensor platforms in a more severe desert condition/environment.

2.2.2 Soil Type

Soil samples were collected at the YPG UXO Standardized Test Site by ERDC to characterize the shallow subsurface (< 3 m). Both surface grab samples and continuous soil borings were acquired. The soils were subjected to several laboratory analyses, including sieve/hydrometer, water content, magnetic susceptibility, dielectric permittivity, X-ray diffraction, and visual description.

Two soil complexes are present within the site, Riverbend-Carrizo and Cristobal-Gunsight. The Riverbend-Carrizo complex is composed of mixed stream alluvium, whereas the Cristobal-Gunsight complex is derived from fan alluvium. The Cristobal-Gunsight complex covers the majority of the site. Most of the soil samples were classified as either a sandy loam or loamy sand, with most samples containing gravel-size particles. All samples had a measured water content less than 7 percent, except for two that contained 11-percent moisture. The majority of soil samples had water content between 1 and 2 percent. Samples containing more than 3 percent were generally deeper than 1 meter.

An X-ray diffraction analysis on four soil samples indicated a basic mineralogy of quartz, calcite, mica, feldspar, magnetite, and some clay. The presence of magnetite imparted a moderate magnetic susceptibility, with volume susceptibilities generally greater than 100 by 105 SI.

For more details concerning the soil properties at the YPG test site, go to www.uxotestsites.org on the Web to view the entire soils description report.

2.2.3 Test Areas

A description of the test site areas at YPG is included in Table 2.

TABLE 2. TEST SITE AREAS

Area	Description
Calibration grid	Contains the 15 standard ordnance items buried in six positions at various angles and depths to allow demonstrator equipment calibration.
Blind grid	Contains 400 grid cells in a 0.16-hectare (0.39-acre) site. The center of each grid cell contains ordnance, clutter, or nothing.

SECTION 3. FIELD DATA

3.1 DATE OF FIELD ACTIVITIES (12 through 14, 16 June 2006)

3.2 AREAS TESTED/NUMBER OF HOURS

Areas tested and total number of hours operated at each site are summarized in Table 3.

TABLE 3. AREAS TESTED AND NUMBER OF HOURS

Area	Number of Hours
Calibration lanes	1.10
Open field	23.80

3.3 TEST CONDITIONS

3.3.1 Weather Conditions

A YPG weather station located approximately 1 mile west of the test site was used to record average temperature and precipitation on a half-hour basis for each day of operation. The temperatures listed in Table 4 represent the average temperature during field operations from 0700 to 1700 hours, while precipitation data represent a daily total amount of rainfall. Hourly weather logs used to generate this summary are provided in Appendix B.

TABLE 4. TEMPERATURE/PRECIPITATION DATA SUMMARY

Date, 2006	Average Temperature, ^{o}F	Total Daily Precipitation, in.
12 June	94.96	0.00
13 June	102.30	0.00
14 June	93.96	0.00
16 June	99.96	0.00

3.3.2 Field Conditions

Geocenter SAIC experienced a dry field and hot weather throughout the survey.

3.3.3 Soil Moisture

Three soil probes were placed at various locations within the site to capture soil moisture data: blind grid, calibration, mogul, and wooded areas. Measurements were collected in percent moisture and were taken twice daily (morning and afternoon) from five different soil depths (1 to 6 in., 6 to 12 in., 12 to 24 in., 24 to 36 in., and 36 to 48 in.) from each probe. Soil moisture logs are included in Appendix C.

3.4 FIELD ACTIVITIES

3.4.1 Setup/Mobilization

These activities included initial mobilization and daily equipment preparation and breakdown. A five-person crew took 2 hours and 14 minutes to perform the initial setup and mobilization. There was 2 hours and 9 minutes of daily equipment preparation and end of the day equipment breakdown lasted 1 hour and 35 minutes.

3.4.2 Calibration

Geocenters SAIC spent a total of 1 hour and 6 minutes in the calibration lanes, of which 31 minutes was spent collecting data.

3.4.3 Downtime Occasions

Occasions of downtime are grouped into five categories: equipment/data checks or equipment maintenance, equipment failure and repair, weather, demonstration site issues, or breaks/lunch. All downtime is included for the purposes of calculating labor costs (section 5) except for downtime due to demonstration site issues. Demonstration site issues, while noted in the daily log, are considered non-chargeable downtime for the purposes of calculating labor costs and are not discussed. Breaks and lunches are discussed in this section and billed to the total site survey area.

3.4.3.1 Equipment/data checks, maintenance. Equipment data checks and maintenance activities accounted for 2 hours and 38 minutes of site usage time. These activities included changing out batteries and performing routine data checks to ensure the data were being properly recorded/collected. Geocenters SAIC spent an additional 2 hours and 23 minutes for breaks and lunches.

3.4.3.2 Equipment failure or repair. 1 hour and 55 minutes was needed to resolve equipment failures that occurred while surveying the open field. The vehicle quit and was later started while troubleshooting. SAIC also replaced a faulty magnetometer and thought they had a hole in the fuel line, which turned out to be just overflow.

3.4.3.3 Weather. No weather delays occurred during the survey.

3.4.4 Data Collection

Geocenters SAIC spent a total time of 23 hours and 48 minutes in the open field area, of which 13 hours and 8 minutes was spent collecting data.

3.4.5 <u>Demobilization</u>

The Geocenters SAIC survey crew went on to conduct a full demonstration of the site. Therefore, demobilization did not occur until 16 June 2006. On that day, it took the crew 7 hours and 39 minutes to break down and pack up their equipment.

3.5 PROCESSING TIME

Geocenters SAIC submitted the raw data from the demonstration activities on the last day of the demonstration, as required. The scoring submittal data were also provided within the required 30-day time frame.

3.6 DEMONSTRATOR’S FIELD PERSONNEL

Rob Seigel
2 field support personnel

3.7 DEMONSTRATOR’S FIELD SURVEYING METHOD

Geocenters SAIC surveyed the open field in a linear manner and in a south-to-north and east-to-west direction, using the width of the array for line spacing.

3.8 SUMMARY OF DAILY LOGS

Daily logs capture all field activities during this demonstration and are located in Appendix D. Activities pertinent to this specific demonstration are indicated in highlighted text.

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 ROC CURVES USING ALL ORDNANCE CATEGORIES

The probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive are shown in Figure 2. Both probabilities plotted against their respective background alarm rate are shown in Figure 3. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

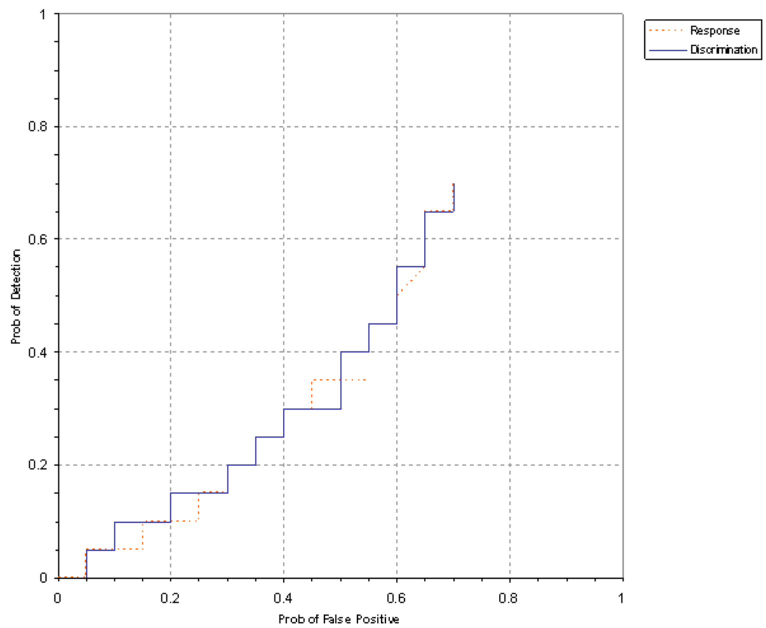

Figure 2. VSEMS/towed probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined.

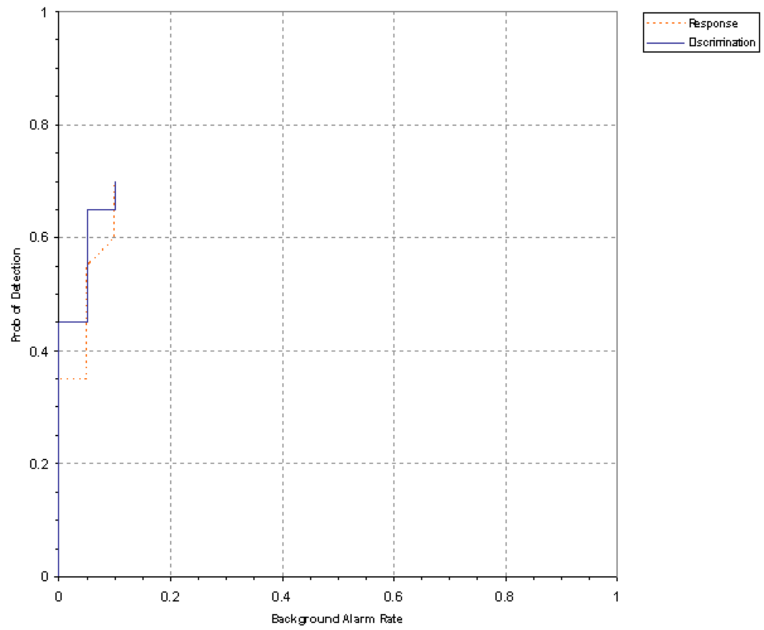

Figure 3. VSEMS/towed probability of detection for response and discrimination stages versus their respective background alarm rate over all ordnance categories combined.

4.2 ROC CURVES USING ORDNANCE LARGER THAN 20 MM

The probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive when only targets larger than 20 mm are scored are shown in Figure 4. Both probabilities plotted against their respective background alarm rate is shown in Figure 5. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

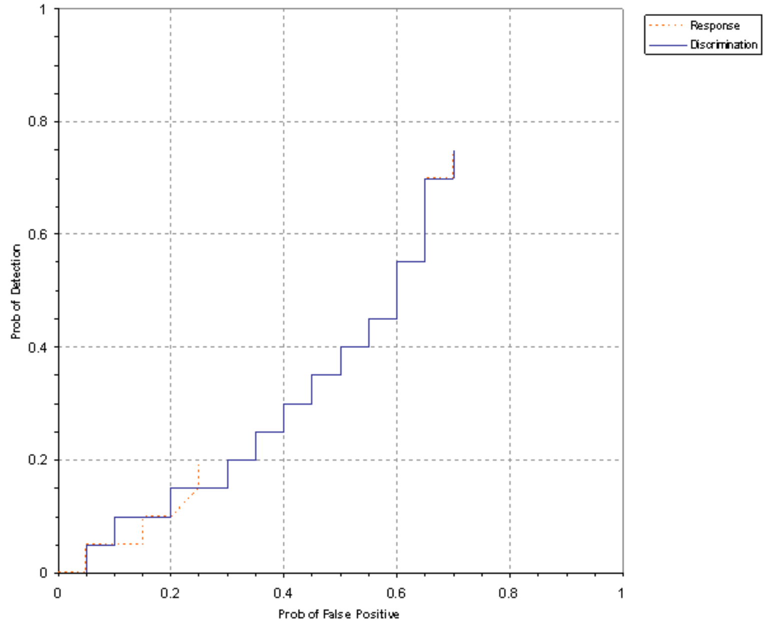

Figure 4. VSEMS/towed probability of detection for response and discrimination stages versus their respective probability of false positive for all ordnance larger than 20 mm.

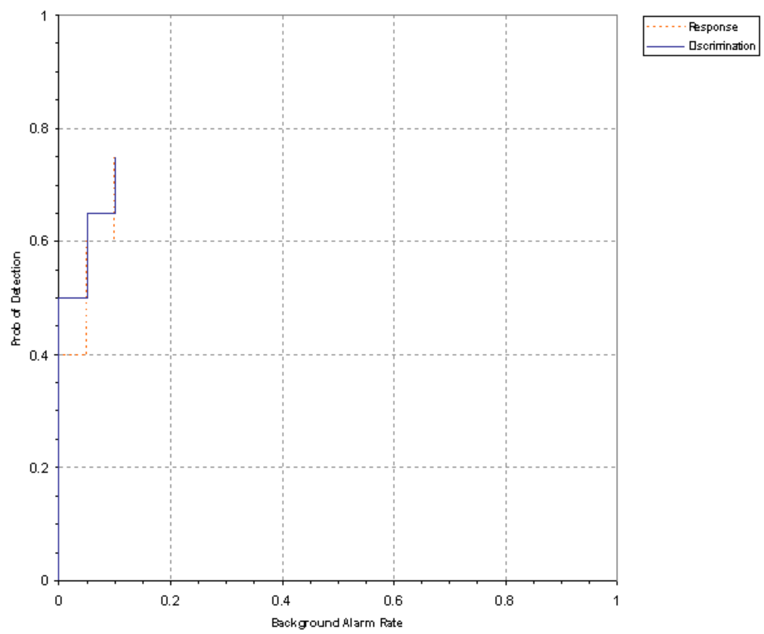

Figure 5. VSEMS/towed probability of detection for response and discrimination stages versus their respective background alarm rate for all ordnance larger than 20 mm.

4.3 PERFORMANCE SUMMARIES

Results for the open field, broken out by size, depth, and nonstandard ordnance, are presented in Table 5 (for cost results, see section 5). Results by size and depth include both standard and nonstandard ordnance. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range (see app A for size definitions). The results are relative to the number of ordnance items emplaced. Depth is measured from the geometric center of anomalies.

The RESPONSE STAGE results are derived from the list of anomalies above the demonstrator-provided noise level. The results for the DISCRIMINATION STAGE are derived from the demonstrator's recommended threshold for optimizing UXO field cleanup by minimizing false digs and maximizing ordnance recovery. The lower 90-percent confidence limit on probability of detection and probability of false positive was calculated assuming that the number of detections and false positives are binomially distributed random variables. All results in Table 5 have been rounded to protect the ground truth. However, lower confidence limits were calculated using actual results.

TABLE 5. SUMMARY OF OPEN FIELD RESULTS FOR VSEMS/TOWED

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.70	0.65	0.75	0.60	0.70	0.90	0.70	0.75	0.55
P_d Low 90% Conf	0.67	0.62	0.70	0.57	0.66	0.85	0.64	0.70	0.45
P_d Upper 90% Conf	0.72	0.69	0.79	0.65	0.75	0.94	0.71	0.80	0.67
P_{fp}	0.70	-	-	-	-	-	0.65	0.80	0.50
P_{fp} Low 90% Conf	0.67	-	-	-	-	-	0.63	0.75	0.27
P_{fp} Upper 90% Conf	0.70	-	-	-	-	-	0.67	0.81	0.73
BAR	0.10	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.65	0.60	0.70	0.50	0.70	0.90	0.60	0.70	0.55
P_d Low 90% Conf	0.61	0.55	0.68	0.47	0.66	0.85	0.58	0.66	0.45
P_d Upper 90% Conf	0.67	0.62	0.76	0.55	0.75	0.94	0.65	0.76	0.67
P_{fp}	0.65	-	-	-	-	-	0.60	0.75	0.40
P_{fp} Low 90% Conf	0.64	-	-	-	-	-	0.60	0.74	0.19
P_{fp} Upper 90% Conf	0.68	-	-	-	-	-	0.64	0.80	0.65
BAR	0.10	-	-	-	-	-	-	-	-

Response Stage Noise Level: 7.00.
Recommended Discrimination Stage Threshold: 233.50.

Note: The recommended discrimination stage threshold values are provided by the demonstrator.

4.4 EFFICIENCY, REJECTION RATES, AND TYPE CLASSIFICATION

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are reported in Table 6.

TABLE 6. EFFICIENCY AND REJECTION RATES

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	0.92	0.04	0.29
With No Loss of P_d	1.00	0.00	0.00

At the demonstrator's recommended setting, the ordnance items that were detected and correctly discriminated were further scored on whether their correct type could be identified (table 7). Correct type examples include 20-mm projectile, 105-mm HEAT Projectile, and 2.75-inch Rocket. A list of the standard type declaration required for each ordnance item was provided to demonstrators prior to testing. For example, the standard type for the three example items are 20 mmP, 105 H, and 2.75 in., respectively.

TABLE 7. CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS UXO

Size	Percentage Correct
Small	0.0
Medium	0.0
Large	0.0
Overall	0.0

Note: The demonstrator did not attempt to provide type classification.

4.5 LOCATION ACCURACY

The mean location error and standard deviations appear in Table 8. These calculations are based on average missed depth for ordnance correctly identified in the discrimination stage. Depths are measured from the closest point of the ordnance to the surface. For the blind grid, only depth errors are calculated because (X, Y) positions are known to be the centers of each grid square.

TABLE 8. MEAN LOCATION ERROR AND STANDARD DEVIATION (M)

	Mean	Standard Deviation
Northing	-0.023	0.165
Easting	-0.032	0.167
Depth	0.138	0.243

SECTION 5. ON-SITE LABOR COSTS

A standardized estimate for labor costs associated with this effort was calculated as follows: the first person at the test site was designated supervisor, the second person was designated data analyst, and the third and following personnel were considered field support. Standardized hourly labor rates were charged by title: supervisor at $95.00/hour, data analyst at $57.00/hour, and field support at $28.50/hour.

Government representatives monitored on-site activity. All on-site activities were grouped into one of ten categories: initial setup/mobilization, daily setup/stop, calibration, collecting data, downtime due to break/lunch, downtime due to equipment failure, downtime due to equipment/data checks or maintenance, downtime due to weather, downtime due to demonstration site issue, or demobilization. See Appendix D for the daily activity log. See section 3.4 for a summary of field activities.

The standardized cost estimate associated with the labor needed to perform the field activities is presented in Table 9. Note that calibration time includes time spent in the calibration lanes as well as field calibrations. Site survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime due to equipment/data checks or maintenance, downtime due to failure, and downtime due to weather.

TABLE 9. ON-SITE LABOR COSTS

	No. People	Hourly Wage	Hours	Cost
Initial Setup				
Supervisor	1	$95.00	2.23	$211.85
Data analyst	1	57.00	2.23	127.11
Field support	3	28.50	2.23	190.67
Subtotal				**$529.63**
Calibration				
Supervisor	1	$95.00	1.10	$104.50
Data analyst	1	57.00	1.10	62.70
Field support	1	28.50	1.10	31.35
Subtotal				**$198.55**
Site Survey				
Supervisor	1	$95.00	23.80	$2261.00
Data analyst	1	57.00	23.80	1356.60
Field support	0	28.50	23.80	0.00
Subtotal				**$3617.60**

See notes at end of table.

TABLE 9 (CONT'D)

	No. People	Hourly Wage	Hours	Cost
Demobilization				
Supervisor	1	$95.00	7.65	$726.75
Data analyst	1	57.00	7.65	436.05
Field support	1	28.50	7.65	218.03
Subtotal				$1380.83
Total				**$5726.61**

Notes: Calibration time includes time spent in the calibration lanes as well as calibration before each data run.

Site survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime due to system maintenance, failure, and weather.

SECTION 6. COMPARISON OF RESULTS TO BLIND GRID DEMONSTRATION (BASED ON COMBINED EM/MAG DATA SETS)

6.1 SUMMARY OF RESULTS FROM BLIND GRID DEMONSTRATION

Table 10 shows the results from the blind grid survey conducted prior to surveying the open field during the same site visit in April 2006. Due to the system utilizing magnetometer type sensors, all results presented in the following section have been based on performance scoring against the ferrous only ground truth anomalies. For more details on the blind grid survey results reference section 2.1.6.

TABLE 10. SUMMARY OF BLIND GRID RESULTS FOR THE VSEMS/TOWED

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.85	0.80	0.85	0.85	0.80	0.95	0.95	0.80	0.30
P_d Low 90% Conf	0.77	0.73	0.74	0.74	0.63	0.75	0.85	0.69	0.08
P_d Upper 90% Conf	0.89	0.89	0.94	0.91	0.89	0.99	0.98	0.91	0.60
P_{fp}	0.90	-	-	-	-	-	0.90	0.95	N/A
P_{fp} Low 90% Conf	0.88	-	-	-	-	-	0.86	0.83	0.00
P_{fp} Upper 90% Conf	0.95	-	-	-	-	-	0.95	0.98	1.00
P_{ba}	0.00	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.85	0.80	0.85	0.85	0.80	0.95	0.95	0.80	0.30
P_d Low 90% Conf	0.77	0.73	0.74	0.74	0.63	0.75	0.85	0.69	0.08
P_d Upper 90% Conf	0.89	0.89	0.94	0.91	0.89	0.99	0.98	0.91	0.60
P_{fp}	0.90	-	-	-	-	-	0.90	0.95	N/A
P_{fp} Low 90% Conf	0.88	-	-	-	-	-	0.86	0.83	0.00
Pfp Upper 90% Conf	0.95	-	-	-	-	-	0.95	0.98	1.00
P_{ba}	0.00	-	-	-	-	-	-	-	-

6.2 COMPARISON OF ROC CURVES USING ALL ORDNANCE CATEGORIES

The P_d^{res} versus the respective P_{fp} over all ordnance categories is shown in Figure 6. The P_d^{disc} versus the respective P_{fp} over all ordnance categories is shown in Figure 7. Horizontal lines illustrate the performance of the demonstrator at the recommended discrimination threshold levels, defining the subset of targets the demonstrator would recommend digging based on discrimination. The ROC curves in this section are a sole reflection of the ferrous only survey.

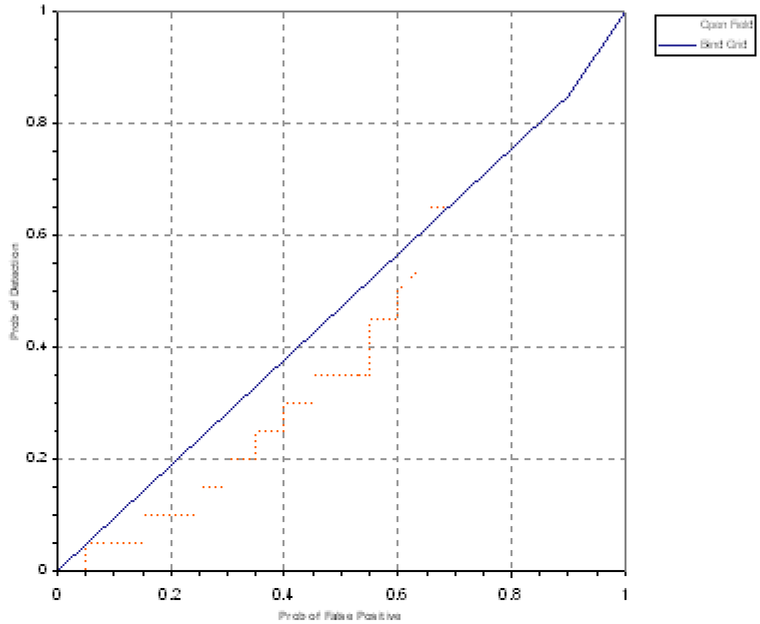

Figure 14. VSEMS dual/towed P_d^{res} stages versus the respective P_{fp} over all ordnance categories combined.

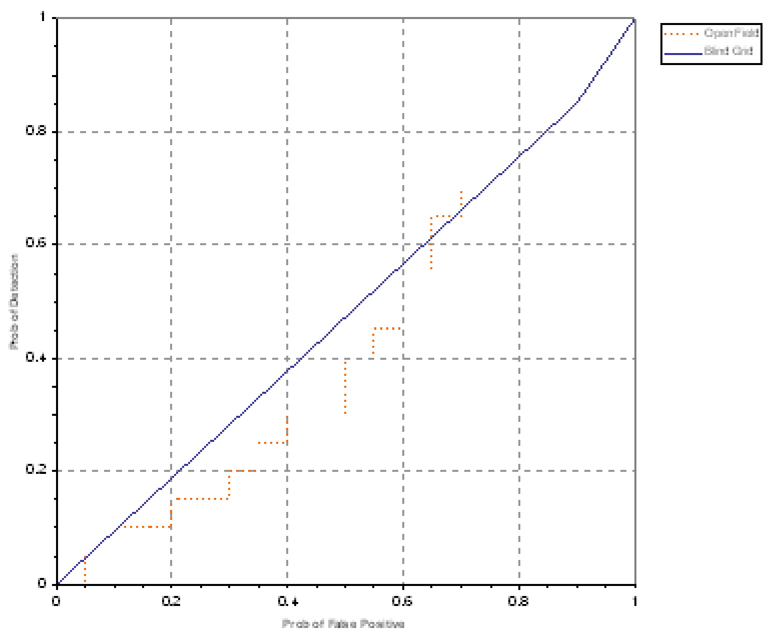

Figure 15. VSEMS/towed P_d^{disc} versus the respective P_{fp} over all ordnance categories combined.

6.3 COMPARISON OF ROC CURVES USING ORDNANCE LARGER THAN 20 MM

The P_d^{res} versus the respective probability of P_{fp} over ordnance larger than 20 mm is shown in Figure 8. The P_d^{disc} versus the respective P_{fp} over ordnance larger than 20 mm is shown in Figure 9. Horizontal lines illustrate the performance of the demonstrator at the recommended discrimination threshold levels, defining the subset of targets the demonstrator would recommend digging based on discrimination.

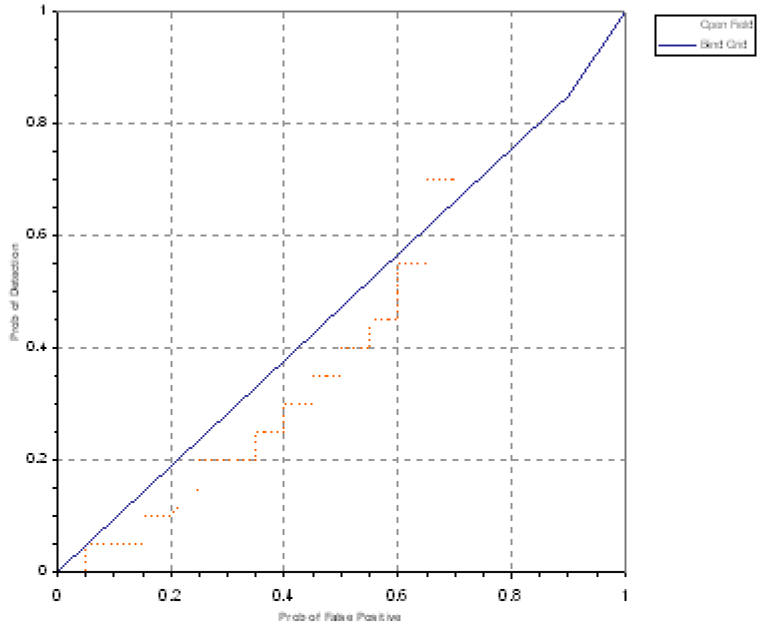

Figure 16. VSEMS/towed P_d^{res} versus the respective P_{fp} for ordnance larger than 20 mm.

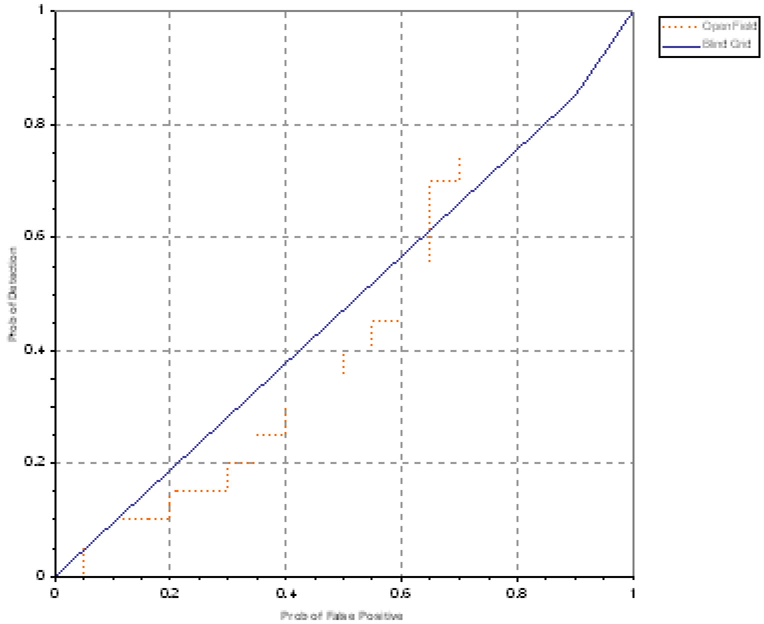

Figure 17. VSEMS/towed P_d^{disc} versus the respective P_{fp} for ordnance larger than 20 mm.

6.4 STATISTICAL COMPARISONS

Statistical Chi-square significance tests were used to compare results between the blind grid and open field scenarios. The intent of the comparison is to determine if the feature introduced in each scenario has a degrading effect on the performance of the sensor system. However, any modifications in the UXO sensor system during the test, like changes in the processing or changes in the selection of the operating threshold, will also contribute to performance differences.

The Chi-square test for comparison between ratios was used at a significance level of 0.05 to compare blind grid to open field with regard to P_d^{res}, P_d^{disc}, P_{fp}^{res} and P_{fp}^{disc}, Efficiency and Rejection Rate. These results are presented in Table 11. A detailed explanation and example of the Chi-square application is located in Appendix A.

TABLE 11. CHI-SQUARE RESULTS - BLIND GRID VERSUS OPEN FIELD

Metric	Small	Medium	Large	Overall
P_d^{res}	Significant	Not Significant	Not Significant	Significant
P_d^{disc}	Significant	Not Significant	Not Significant	Significant
P_{fp}^{res}	-	-	-	Significant
P_{fp}^{disc}	-	-	-	Significant
Efficiency	-	-	-	Significant
Rejection rate	-	-	-	Not Significant

SECTION 7. APPENDIXES

APPENDIX A. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced ordnance item.

Detection: An anomaly location that is within R_{halo} of an emplaced ordnance item.

Munitions and Explosives Of Concern (MEC): Specific categories of military munitions that may pose unique explosive safety risks, including UXO as defined in 10 USC 101(e)(5), DMM as defined in 10 USC 2710(e)(2) and/or munitions constituents (e.g. TNT, RDX) as defined in 10 USC 2710(e)(3) that are present in high enough concentrations to pose an explosive hazard.

Emplaced Ordnance: An ordnance item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., non-ordnance item) buried by the government at a specified location in the test site.

R_{halo}: A pre-determined radius about the periphery of an emplaced item (clutter or ordnance) within which a location identified by the demonstrator as being of interest is considered to be a response from that item. If multiple declarations lie within R_{halo} of any item (clutter or ordnance), the declaration with the highest signal output within the R_{halo} will be utilized. For the purpose of this program, a circular halo 0.5 meters in radius will be placed around the center of the object for all clutter and ordnance items less than 0.6 meters in length. When ordnance items are longer than 0.6 meters, the halo becomes an ellipse where the minor axis remains 1 meter and the major axis is equal to the length of the ordnance plus 1 meter.

Small Ordnance: Caliber of ordnance less than or equal to 40 mm (includes 20-mm projectile, 40-mm projectile, submunitions BLU-26, BLU-63, and M42).

Medium Ordnance: Caliber of ordnance greater than 40 mm and less than or equal to 81 mm (includes 57-mm projectile, 60-mm mortar, 2.75 in. Rocket, MK118 Rockeye, 81-mm mortar).

Large Ordnance: Caliber of ordnance greater than 81 mm (includes 105-mm HEAT, 105-mm projectile, 155-mm projectile, 500-pound bomb).

Shallow: Items buried less than 0.3 meter below ground surface.

Medium: Items buried greater than or equal to 0.3 meter and less than 1 meter below ground surface.

Deep: Items buried greater than or equal to 1 meter below ground surface.

Response Stage Noise Level: The level that represents the point below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the blind grid test area.

Discrimination Stage Threshold: The demonstrator selected threshold level that they believe provides optimum performance of the system by retaining all detectable ordnance and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for n independent trials with the probability p of success and the probability 1-p of failure being the same for each trial. The number of successes x observed in the n trials is an estimate of p and is considered to be a binomially distributed random variable.

RESPONSE AND DISCRIMINATION STAGE DATA

The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}) and those that do not correspond to any known item, termed background alarms.

The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the RESPONSE STAGE, the demonstrator provides the scoring committee with the location and signal strength of all anomalies that the demonstrator has deemed sufficient to warrant further investigation and/or processing as potential emplaced ordnance items. This list is generated with minimal processing (e.g., this list will include all signals above the system noise threshold). As such, it represents the most inclusive list of anomalies.

The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such, and to reject clutter. For the same locations as in the RESPONSE STAGE anomaly list, the DISCRIMINATION STAGE list contains the output of the algorithms applied in the discrimination-stage processing. This list is prioritized based on the demonstrator's determination that an anomaly location is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For electronic signal processing, priority ranking is based on algorithm output. For other systems, priority ranking is based on human judgment. The demonstrator also selects the threshold that the demonstrator believes will provide "optimum" system performance, (i.e., that retains all the detected ordnance and rejects the maximum amount of clutter).

Note: The two lists provided by the demonstrator contain identical numbers of potential target locations. They differ only in the priority ranking of the declarations.

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/(No. of emplaced ordnance in the test site).

Response Stage False Positive (fp^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of False Positive (P_{fp}^{res}): P_{fp}^{res} = (No. of response-stage false positives)/(No. of emplaced clutter items).

Response Stage Background Alarm (ba^{res}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind Grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open Field only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{fp}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can therefore be written as $P_d^{res}(t^{res})$, $P_{fp}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to response-stage data that discriminates ordnance from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to ordnance, as well as those that the demonstrator has high confidence correspond to non-ordnance or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced ordnance in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm (ba^{disc}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can therefore be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] Figure A-1 shows how P_d versus P_{fp} and P_d versus BAR are combined into ROC curves. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

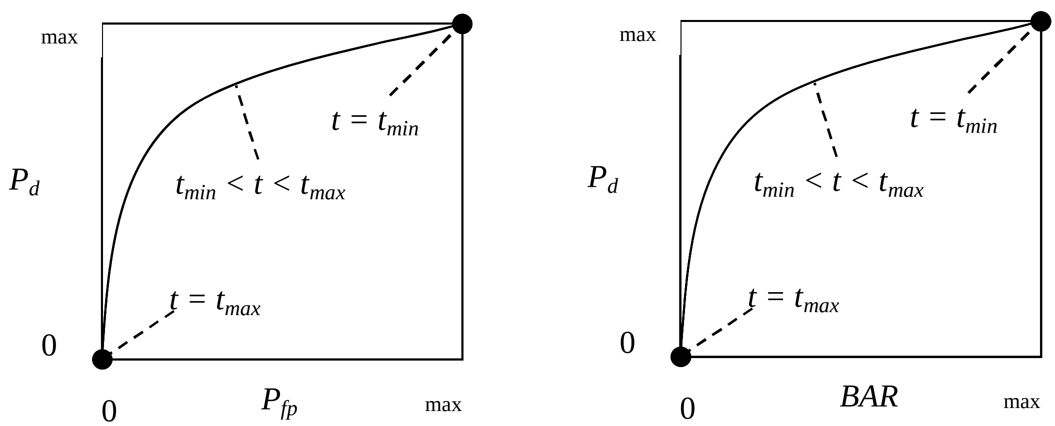

Figure A-1. ROC curves for open field testing. Each curve applies to both the response and discrimination stages.

---

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a pre-determined and fixed number of detection opportunities (some of the opportunities are located over ordnance and others are located over clutter or blank spots). In an open field scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open field ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the blind grid test sites are true ROC curves.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. The efficiency measures the amount of detected ordnance retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$; Measures (at a threshold of interest), the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage t_{min}) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the ordnance initially detected in the response stage was retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{fp}^{res}(t_{min}^{res})]$; Measures (at a threshold of interest), the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage tmin). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$.
Open field: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})]$).

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON EXPLANATION:

The Chi-square test for differences in probabilities (or 2 x 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations (ref 3).

A 2 x 2 contingency table is used in the Standardized UXO Technology Demonstration Site Program to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X’s system is significantly degraded by the more challenging terrain feature introduced. The test statistic of the 2 x 2 contingency table is the

Chi-square distribution with one degree of freedom. Since an association between the more challenging terrain feature and relatively degraded performance is sought, a one-sided test is performed. A significance level of 0.05 is chosen which sets a critical decision limit of 2.71 from the Chi-square distribution with one degree of freedom. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's test is used and the critical decision limit for one-sided tests is the chosen significance level, which in this case is 0.05. With Fischer's test, if the test statistic is less than the critical value, the proportions are considered to be significantly different.

Standardized UXO Technology Demonstration Site examples, where blind grid results are compared to those from the open field and open field results are compared to those from one of the scenarios, follow. It should be noted that a significant result does not prove a cause and effect relationship exists between the two populations of interest; however, it does serve as a tool to indicate that one data set has experienced a degradation in system performance at a large enough level than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying each of the three progressively more difficult areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind grid	Open field	Moguls
P_d^{res}	100/100 = 1.0	8/10 = .80	20/33 = .61
P_d^{disc}	80/100 = 0.80	6/10 = .60	8/33 = .24

P_d^{res}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the response stage, all 100 ordnance out of 100 emplaced ordnance items were detected in the blind grid while 8 ordnance out of 10 emplaced were detected in the open field. Fischer's test must be used since a 100 percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open field relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 emplaced ordnance items were correctly discriminated as ordnance in blind grid testing while 6 ordnance out of 10 emplaced were correctly discriminated as such in open field-testing. Those four values are used to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 2.71, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the response stage, 8 out of 10 and 20 out of 33 are used to calculate a test statistic of 0.56. Since the test statistic is less than the critical value of 2.71, the two response stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{disc}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the discrimination stage, 6 out of 10 and 8 out of 33 are used to calculate a test statistic of 2.98. Since the test statistic is greater than the critical value of 2.71, the smaller discrimination stage detection rate is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the ability of demonstrator X to correctly discriminate seems to have been degraded by the mogul terrain relative to results from the flat open field using the same system.

APPENDIX B. DAILY WEATHER LOGS

Date, 2006	Time	Average Temperature, ˚C	Average Precipitation, in.
6/12	0700	26.1	0.00
	0800	27.9	0.00
	0900	29.3	0.00
	1000	31.1	0.00
	1100	33.1	0.00
	1200	34.8	0.00
	1300	36.5	0.00
	1400	37.2	0.00
	1500	38.9	0.00
	1600	39.5	0.00
	1700	39.3	0.00
6/13	0700	28.4	0.00
	0800	32.5	0.00
	0900	34.7	0.00
	1000	37.0	0.00
	1100	38.3	0.00
	1200	39.3	0.00
	1300	40.2	0.00
	1400	40.6	0.00
	1500	41.7	0.00
	1600	42.3	0.00
	1700	41.8	0.00
6/14	0700	24.8	0.00
	0800	27.0	0.00
	0900	29.9	0.00
	1000	31.3	0.00
	1100	32.9	0.00
	1200	34.4	0.00
	1300	36.1	0.00
	1400	36.9	0.00
	1500	37.5	0.00
	1600	38.0	0.00
	1700	39.1	0.00

Date, 2006	Time	Average Temperature, ˚C	Average Precipitation, in.
6/16	0700	27.0	0.00
	0800	31.3	0.00
	0900	33.8	0.00
	1000	35.0	0.00
	1100	36.7	0.00
	1200	37.4	0.00
	1300	38.4	0.00
	1400	39.3	0.00
	1500	40.0	0.00
	1600	40.2	0.00
	1700	40.1	0.00

APPENDIX C. SOIL MOISTURE

Date: 6/12/2006			
Times: 0830 through 1300			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration lanes	0 to 6	1.8	1.6
	6 to 12	2.3	2.2
	12 to 24	3.7	3.8
	24 to 36	3.7	3.7
	36 to 48	4.3	4.2
Moguls	0 to 6	1.8	1.8
	6 to 12	6.5	6.4
	12 to 24	3.8	3.8
	24 to 36	4.9	4.9
	36 to 48	6.6	6.4
Desert extreme	0 to 6	5.1	4.9
	6 to 12	3.8	3.8
	12 to 24	3.2	3.0
	24 to 36	4.1	4.0
	36 to 48	4.0	4.0

Date: 6/13/2006			
Times: 0630 through 1330			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration lanes	0 to 6	1.7	1.5
	6 to 12	2.1	2.1
	12 to 24	3.8	3.8
	24 to 36	3.8	3.8
	36 to 48	4.3	4.2
Moguls	0 to 6	1.9	1.8
	6 to 12	6.4	6.4
	12 to 24	3.8	3.8
	24 to 36	4.9	4.9
	36 to 48	6.8	6.4
Desert extreme	0 to 6	6.7	6.4
	6 to 12	3.8	3.8
	12 to 24	3.3	3.3
	24 to 36	4.1	4.0
	36 to 48	4.1	4.1

Date: 6/14/2006			
Times: 0600 through 1330			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration lanes	0 to 6	1.4	1.6
	6 to 12	2.1	2.1
	12 to 24	3.8	3.8
	24 to 36	3.8	3.8
	36 to 48	4.3	4.2
Moguls	0 to 6	1.7	1.6
	6 to 12	6.5	6.3
	12 to 24	3.8	3.8
	24 to 36	4.9	4.9
	36 to 48	5.7	5.5
Desert extreme	0 to 6	9.7	9.1
	6 to 12	3.8	3.8
	12 to 24	3.3	3.3
	24 to 36	4.1	4.0
	36 to 48	4.1	4.1

Date: 6/16/2006			
Times: 0630 through 1230			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration lanes	0 to 6	1.5	1.7
	6 to 12	2.2	2.1
	12 to 24	3.8	3.7
	24 to 36	3.8	3.8
	36 to 48	4.3	4.3
Moguls	0 to 6	1.8	1.8
	6 to 12	3.8	3.8
	12 to 24	3.8	3.8
	24 to 36	4.9	4.9
	36 to 48	5.9	5.7
Desert extreme	0 to 6	3.8	3.8
	6 to 12	3.8	3.8
	12 to 24	3.2	3.2
	24 to 36	4.1	4.1
	36 to 48	4.1	4.1

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Pattern	Field Conditions	
06/12/2006	5	CALIBRATION LANES	735	949	134	INITIAL SETUP	Set up test equipment. Unit vehicle-towed array - VSEMS.	NA	NA	Sunny	Warm
06/12/2006	3	CALIBRATION LANES	949	1020	31	COLLECTING DATA	Ran calibration grid south to north, west to east; completed.	GPS	Linear	Sunny	Hot
06/12/2006	3	CALIBRATION LANES	1020	1055	35	DOWNTIME DUE TO EQUIPMENT MAINT/CHECK	Downloaded data from the calibration grid and checked data; data were good.	NA	NA	Sunny	Hot
06/12/2006	3	BLIND TEST GRID	1055	1132	37	COLLECTING DATA	Ran blind grid south to north, west to east; completed.	GPS	Linear	Sunny	Hot
06/12/2006	2	BLIND TEST GRID	1132	1157	25	DOWNTIME DUE TO EQUIPMENT MAINT/CHECK	Downloaded data from the blind grid and checked data; data were good.	NA	NA	Sunny	Hot
06/12/2006	2	BLIND TEST GRID	1157	1249	52	BREAK/LUNCH	Lunch	NA	NA	Sunny	Hot
06/12/2006	2	OPEN FIELD	1249	1500	131	COLLECTING DATA	Ran open field, south to north, east to west, grids D4, D5 to G4.	GPS	Linear	Sunny	Hot
06/12/2006	2	OPEN FIELD	1500	1529	30	DOWNTIME DUE TO EQUIPMENT MAINT/CHECK	Downloaded data from the open field and checked data; data were good.	NA	NA	Sunny	Hot
06/12/2006	2	OPEN FIELD	1529	1622	53	DAILY START, STOP	Breakdown end of day	NA	NA	Sunny	Hot
06/13/2006	2	OPEN FIELD	519	620	61	DAILY START, STOP	Set up test equipment.	NA	NA	Clear	Cool
06/13/2006	2	OPEN FIELD	620	659	39	COLLECTING DATA	Continued to run the open field, grids D3 to G3.	GPS	Linear	Sunny	Cool

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Pattern	Field Conditions	
06/13/2006	2	OPEN FIELD	659	707	8	DOWNTIME DUE TO EQUIPMENT FAILURE	Replaced a faulty magnetometer S/N 75042 with S/N 75038; replacement was good.	NA	NA	Sunny	Warm
06/13/2006	2	OPEN FIELD	707	734	27	BREAK/LUNCH	Break	NA	NA	Sunny	Warm
06/13/2006	2	OPEN FIELD	734	736	2	DOWNTIME DUE TO EQUIPMENT MAINT/CHECK	Ran clear area to null the system after changing out the magnetometer.	GPS	Linear	Sunny	Warm
06/13/2006	2	OPEN FIELD	736	1049	193	COLLECTING DATA	Continued to run the open field, D3 to G3.	GPS	Linear	Sunny	Warm
06/13/2006	2	OPEN FIELD	1049	1123	34	BREAK/LUNCH	Lunch	NA	NA	Sunny	Hot
06/13/2006	2	OPEN FIELD	1123	1233	70	DOWNTIME DUE TO EQUIPMENT MAINT/CHECK	System was hot due to ambient air temperature, separated the GPS receivers to provide airflow to aid in cooling components. Fueled vehicle.	NA	NA	Sunny	Hot
06/13/2006	2	OPEN FIELD	1233	1356	83	COLLECTING DATA	Continued to run the open field; D2 to G2.	GPS	Linear	Sunny	Hot
06/13/2006	2	OPEN FIELD	1356	1501	65	DOWNTIME DUE TO EQUIPMENT FAILURE	Tow vehicle had a hole in the fuel line, vehicle towed from field. Troubleshot fuel system. No hole, fuel leak was actually fuel coming from the overflow.	NA	NA	Sunny	Hot
06/13/2006	2	OPEN FIELD	1501	1521	20	DAILY START, STOP	Breakdown end of day	NA	NA	Sunny	Hot

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Pattern	Field Conditions	
06/14/2006	2	OPEN FIELD	513	621	68	DAILY START, STOP	Set up test equipment. Note: Installed 2,brushless 12-VDC cooling fans to the front and side of the GPS receivers to aid in cooling the components.	NA	NA	Clear	Cool
06/14/2006	2	OPEN FIELD	621	1026	245	COLLECTING DATA	Continued to run the open field, grids A2, A3 to C2, C3.	GPS	Linear	Sunny	Cool
06/14/2006	2	OPEN FIELD	1026	1108	42	DOWNTIME DUE TO EQUIPMENT FAILURE	Tow vehicle quit, troubleshot problem, could not determine; while troubleshooting the vehicle was able to be started, cause unknown.	NA	NA	Sunny	Hot
06/14/2006	2	OPEN FIELD	1108	1112	4	DOWNTIME DUE TO EQUIPMENT MAINT/CHECK	Ran system over to the clear area to null the system.	GPS	Linear	Sunny	Hot
06/14/2006	2	OPEN FIELD	1112	1209	57	COLLECTING DATA	Continued to run the open field A4, A5 to C4, C5.	GPS	Linear	Sunny	Hot
06/14/2006	2	OPEN FIELD	1209	1247	38	DOWNTIME DUE TO EQUIPMENT MAINT/CHECK	Downloaded data from the open field and checked data; data were good.	NA	NA	Sunny	Hot
06/14/2006	2	OPEN FIELD	1247	1409	82	BREAK/LUNCH	Lunch Note: Awaiting other crew (MSEMS) to finish the blind grid.	NA	NA	Sunny	Hot
06/14/2006	2	OPEN FIELD	1409	1449	40	COLLECTING DATA	Ran blind grid west to east; south to north; completed.	GPS	Linear	Sunny	Hot

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min	Operational Status	Operational Status Comments	Track Method	Pattern	Field Conditions	
06/14/2006	2	OPEN FIELD	1449	1503	14	DOWNTIME DUE TO EQUIPMENT MAINT/CHECK	Downloaded data from the grid field.	NA	NA	Sunny	Hot
06/14/2006	2	OPEN FIELD	1503	1525	22	DAILY START, STOP	Breakdown end of day	NA	NA	Sunny	Hot
06/16/2006	3	OPEN FIELD	621	656	35	DEMOBILIZATION	Disassembled the VSEMS system.	NA	NA	Clear	Cool
06/16/2006	2	OPEN FIELD	656	1400	424	DEMOBILIZATION	Disassembled the VSEMS system.	NA	NA	Clear	Warm

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

APPENDIX E.   REFERENCES

1. Standardized UXO Technology Demonstration Site Handbook, DTC Project No. 8-CO-160-000-473, Report No. ATC-8349, March 2002.

2. Aberdeen Proving Ground Soil Survey Report, October 1998.

3. Data Summary, UXO Standardized Test Site: APG Soils Description, May 2002.

4. Yuma Proving Ground Soil Survey Report, May 2003.

APPENDIX F. ABBREVIATIONS

ADST	=	Aberdeen Data Services Team
AOI	=	area of interest
APG	=	Aberdeen Proving Ground
ATC	=	U.S. Army Aberdeen Test Center
ATSS	=	Aberdeen Test and Support Services
BAH	=	Booz Allen Hamilton
CEHNC	=	Corps of Engineers - Huntsville Center
DMM	=	discarded military munitions
E	=	efficiency
EM	=	electromagnetic
EMI	=	electromagnetic interference
ERDC	=	U.S. Army Corps of Engineers Engineering Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
EQT	=	Army Environmental Quality Technology Program
GPS	=	Global Positioning System
HEAT	=	high-explosive antitank
JPG	=	Jefferson Proving Ground
M	=	standard deviation
MAG	=	magnetometer
MEC	=	munitions and explosives of concern
METDC	=	Military Environmental Technology Demonstration Center
NS	=	nonstandard
POC	=	point of contact
QA	=	quality assurance
QC	=	quality control
ROC	=	receiver-operating characteristic
RTK	=	real time kinematic
SAIC	=	Science Applications International Corporation
SERDP	=	Strategic Environmental Research and Development Program
USAEC	=	U.S. Army Environmental Command
UTM	=	Universal Transverse Mercator
UXO	=	unexploded ordnance
VSEMS	=	vehicular simultaneous EMI and magnetometer system
YPG	=	U.S. Army Yuma Proving Ground

APPENDIX G. DISTRIBUTION LIST

DTC Project No. 8-CO-160-UXO-021

Addressee	No. of Copies
Booz Allen Hamilton ATTN: Mr. Patrick McDonnell 4692 Millennium Drive Suite 200 Belcamp, MD 21017-1535	1
Geocenters SAIC ATTN: Mr. Rob Siegel 7 Wells Avenue Newton, MA 02459	1
Commander U.S. Army Aberdeen Test Center ATTN: TEDT-AT-SLE (Mr. J. Stephen McClung) 400 Colleran Road Aberdeen Proving Ground, MD 21005-5059	1
Defense Technical Information Center 8725 John J. Kingman Road, Suite 0944 Fort Belvoir, VA 22060-6218	PDF

Secondary distribution is controlled by Commander, U.S. Army Environmental Command, ATTN: IMAE-RTA.